Chapter 9

Conclusions

In Section 9.1 we first summarize the major original contributions of this thesis. Finally, in Section 9.2 we point out the limitations of the approach, the problems that still need to be addressed, and some possible solutions.

9.1 Summary of Original Contributions

In the introductory Section 1.3 we tried to assess the state of the art of soil moisture data assimilation. Having described our approach in detail in the previous Chapters, we now summarize the major original contributions of this work. In a nutshell, the five major contributions of this thesis are (1) the implementation of the optimal representer algorithm for a hydrologic data assimilation problem, (2) the development of a land surface model suitable for data assimilation, (3) the formulation of a general downscaling methodology for L-band passive microwave images, (4) the investigation of a number of topics which are crucial to the design of an operational soil moisture data assimilation system, and (5) the detailed assessment of the computational requirements of the algorithm for the soil moisture assimilation problem. Here is a more detailed summary of our contributions:

1. The optimal indirect iterated representer algorithm has been applied for the first time to a hydrologic data assimilation problem.

   (a) The assimilation algorithm is fully four-dimensional. We account explicitly for vertical and horizontal correlations.

   (b) The hydrologic model enters the assimilation as a weak constraint. Model error (or process noise) is taken into account.

   (c) The iterated representer method has been extended to allow for a nonlinear measurement operator, for a nonlinear dependence of the state on the parameters, and for a state-dependent coefficient multiplying the process noise.

   (d) Satellite radiances (or brightness temperatures) are assimilated directly. No off-line inversions of the remote sensing observations are necessary.

   (e) The formulation of the algorithm allows for the assimilation of various other data types, not only L-band passive microwaves.

   (f) Posterior error covariance calculations have been formulated for land surface hydrologic applications.

2. A land surface model suitable for hydrologic data assimilation has been developed.

   (a) Our model captures the key physical processes of the land-atmosphere boundary, and at the same time it is very computationally efficient.

   (b) The computational efficiency has been achieved foremost by dividing the model domain into laterally uncoupled one-dimensional vertical columns, which we call estimation pixels. Horizontal coupling has been incorporated through correlations of the inputs and of the model and parameter errors.

   (c) Our model fulfills the basic differentiability requirement for use in a variational assimilation algorithm.

   (d) For the soil moisture dynamics, any number of vertical nodes can be specified. It is easy to adjust the resolution of the soil moisture profile according to the problem at hand.

   (e) Soil and canopy temperatures are modeled together with soil moisture. This obviates the need to independently specify soil and canopy temperatures for the Radiative Transfer model. It also opens up the possibility of assimilating remotely sensed soil and canopy temperatures.

   (f) We have developed the adjoint of our land surface model.

3. A general downscaling capability has been incorporated into the assimilation algorithm, making it possible to effectively increase the resolution of the remote sensing brightness images, or equivalently, to estimate the land surface states at a scale finer than the resolution of the brightness images (Section 4.7).

4. We have conducted a series of synthetic experiments to test our land surface assimilation algorithm. These experiments demonstrate the usefulness of L-band radiobrightness data for soil moisture estimation under realistic conditions. We have assessed the optimality of the estimates for all experiments.

   (a) For a series of synthetic experiments under ideal and nonideal conditions, the assimilation algorithm can estimate the (synthetic) true land surface states to a high degree of accuracy.

   (b) From our results we conclude that large-scale soil moisture estimation from L-band passive microwave data is feasible.

   (c) In two reference experiments, we have successfully demonstrated the ability of the estimation algorithm to estimate the initial soil moisture conditions and the moisture flux boundary condition at the land surface (Sections 6.1 and 6.2).

   (d) We have investigated two typical downscaling scenarios which combine coarse-scale remote sensing data with fine-scale information from the model inputs (Section 6.3). Even for downscaling ratios of one to sixteen (each observation pixel contains sixteen estimation pixels), the estimate captures many of the fine-scale features of the true fields. This implies that brightness images with a resolution of $50km$ may be used to infer soil moisture on the scale of ten to twenty kilometers, provided sufficiently accurate model inputs are available at the finer scale.

(e) Repeat frequencies of up to three days for brightness images allow for satisfactory estimation of soil moisture conditions (Section 6.4). Generally, the repeat frequency should not be smaller than the typical frequency for rainstorms. If brightness data are available less frequently, the soil moisture estimates deteriorate rapidly.

(f) We have assessed the influence of the length of the data assimilation interval and the problem of reinitializing the windows in an operational fashion (Section 7.1). For an operational setup, it seems best to match the assimilation window approximately to an interstorm period. The assimilation intervals must be chosen such that there is sufficient time for the initial error covariance to evolve before the first observation time. This mitigates the negative effects of naively reinitializing the assimilation windows.

(g) We have demonstrated the ability of the assimilation algorithm to satisfactorily estimate soil moisture even if quantitative precipitation data are not available (Section 7.2).

(h) The assimilation algorithm is also capable of satisfactorily estimating soil moisture even if the soil hydraulic parameters are only poorly known (Section 7.3).

(i) For all experiments, we have thoroughly assessed the optimality of the estimates by examining the posterior data residuals and the value of the reduced objective function.

Even under ideal assimilation conditions, the nonlinear structure of the hydrologic model and the measurement operator leads to deviations of the residuals' sample cumulative distribution function from a normal distribution.

The residuals and the reduced objective function consistently show when the assimilation conditions were not ideal and the estimates were therefore suboptimal.

5. The computational requirements of the assimilation algorithm have been assessed in detail.

(a) For a typical application, the computational effort of the iterated indirect representer method grows less than linearly with the number of scalar data (Section 8.2.1).

(b) The computational effort grows only linearly with the number of pixels (Section 8.2.2).

(c) The prior statistics and the length of the assimilation interval strongly influence the computational requirements (Section 8.1.2 and 8.1.3). Assimilation windows of a few days are computationally very attractive while providing estimates that are close to optimal.

(d) For the land surface data assimilation problem discussed in this thesis, the iterated indirect representer approach is competitive from a computational perspective with suboptimal sequential Monte Carlo methods such as the Ensemble Kalman Filter (EnKF) and error subspace statistical estimation (ESSE) (Section 8.3). For a given computational cost, the trade-off is between the optimality of the estimates (representer method) and the availability of posterior covariance information (EnKF, ESSE).

9.2 Limitations and Problems to be Addressed

We now provide some comments on the research presented throughout this thesis. These comments are meant to serve three main purposes. First, we hope to raise the reader's awareness of any difficulties that may arise when implementing the methods described in this work. Second, we suggest areas that require more research, and perhaps entirely new directions of approaching the problem. Last but not least, we point at possible solutions that we think are appropriate but did not investigate for lack of time.

Given the highly nonlinear structure of land-atmosphere processes, and given the high complexity of real world applications, any large-scale land surface data assimilation algorithm will necessarily be a compromise between realistic physical representations and computational feasibility. A lot more research still needs to be done before an operational soil moisture data assimilation package can be set up. Here is a list of caveats, possible solutions, and suggestions for future research:

1. So far, the assimilation algorithm has only been verified with a few synthetic experiments. A field test and further synthetic experiments must be carried out.

   (a) All results have been obtained with a few synthetic experiments. Even though we made an effort to mimic realistic conditions, we have only discussed a total of three different realizations. There can be no guarantee that the results remain completely unchanged if more synthetic experiments are added which rely on different random seeds.

   (b) The results may depend on the particular choices of the error covariances. As we have seen Section 8.1.2, the computational demand is already quite sensitive to the prior statistics. To corroborate the findings, these intuitive choices will have to be validated in a field application, or, alternatively, the same outcome would have to be found for synthetic experiments with many different setups.

   (c) Before the assimilation algorithm can be applied to any particular field experiment, a robust calibration of the hydrologic model must be carried out for the field site in question.

   (d) The next, indispensable step is to conduct a field test of the algorithm by assimilating the SGP97 ESTAR observations. This will bring up difficult issues of model bias. Using the hypothesis tests described in Sections 2.3.6 and 2.4.1, we will then be able to thoroughly check whether the model and the statistical assumptions are appropriate in the field. In other words, the field test will help to identify realistic prior error statistics for land surface data assimilation.

2. The computation of the posterior error covariances must be implemented.

   (a) Ideally, when running the estimator in an operational mode, the prior error covariance of the initial condition should be specified or at least approximated respecting the posterior error covariances of the previous assimilation window. If this turns out to be too expensive to do operationally, the computation of the posterior error covariances in research studies will still provide valuable insights into the accuracy of the estimates and the operating conditions of the assimilation algorithm.

(b) It is certainly unnecessary in any case to compute the full posterior error covariances of the state and the measurements. Suitable approximations for the calculation of the posterior error covariances must be identified.

(c) A priori data compression should be implemented to decrease the cost of calculating the posterior covariances. There may also be additional savings from a priori data compression that are not realized with the indirect representer method. This topic is worth a detailed investigation because of the practical relevance of such savings.

3. The application and interpretation of Richards' equation at the regional scale clearly deserve more attention.

   (a) In our approach, the second-order Richards' equation is but a means to model moisture fronts that move nonlinearly in both directions. Any simple formulation capable of producing such upward and downward movement will in practice resemble Richards' equation.

   (b) Consequently, we do not claim that the small-scale soil hydraulic properties published in the literature can necessarily be interpreted as the parameters that we should use for our large-scale Richards' equation. Moreover, we do not interpret the soil moisture values in different vertical layers as a soil moisture profile that could be verified with ground-based point observations.

   (c) In our interpretation, Richards' equation can be viewed as governing a nonlinear multilayer reservoir of soil moisture that is reasonably consistent with the true land surface fluxes and with the groundwater recharge. The verification of this hypothesis and of the validity of the profile estimates will have to come from data assimilation experiments which are conducted over long experiment periods in a quasi-operational manner (Section 6.1.3).

4. In the current implementation, the assimilation algorithm depends heavily on the information needed to run the hydrologic model. To relax this dependence, other data types must be assimilated.

   (a) L-band brightness temperatures are determined not only by soil moisture, but also by soil temperature. Therefore our ability to estimate soil moisture from observations of the L-band brightness temperature critically depends on how accurately we know soil temperature. The experiments of Chapters 6 and 7 are not designed to test the capability of the algorithm to estimate the soil temperature. Instead, the soil temperature was rather well known by design. In an operational setting, such good knowledge of soil temperature can only be achieved by assimilating supplemental data, for example infrared remote sensing observations.

   (b) If satellite derived soil skin temperatures are assimilated, the one-layer soil temperature force-restore approximation may have to be extended to a two layer model. This would, however, conflict with the desire to keep the model as computationally efficient as possible for the transition towards continental-scale operational assimilation.

(c) The precipitation inputs to the land surface model are crucial for the quality of the soil moisture estimates. In the current implementation, errors in the micro-meteorologic forcings including precipitation are lumped into the model error terms. In other words, we treat precipitation as a model parameter which is not explicitly estimated.

If the precipitation data are poor, it may be beneficial to assimilate these data into a coupled land surface and precipitation model. Note, however, that a simple scheme would not offer much information on the complicated nature of the precipitation processes. A sophisticated precipitation model, on the other hand, may prove hard to invert or too computationally expensive.

5. The experiments of Chapters 6 and 7 represent only a small fraction of the possibilities to study land surface data assimilation problems with our algorithm. Many other useful synthetic experiments can be thought up.

   (a) Together with the soil properties, the land cover parameters are critical for estimating soil moisture and other land surface variables. Therefore, the impact of the vegetation on the quality of the estimates must be investigated in more detail.

   (b) We have not fully explored the potential to estimate land surface variables other than soil moisture. Suitable experiments must be designed to assess the capability of the algorithm to infer soil temperature, canopy temperature, and the land surface fluxes from L-band passive microwaves and possibly other remote sensing observations.

   (c) Operational brightness observations at higher frequencies are already available or will soon become available. The algorithm could be used to investigate the trade-off of assimilating many data of poorer quality with respect to soil moisture, such as C-band ($5.3 GHz$) observations, versus fewer data of higher quality, such as L-band ($1.4 GHz$) observations.

6. The land surface model may require modifications for certain applications.

   (a) The measurement operator for the general downscaling methodology must be re-evaluated when only C-band brightness temperatures are available for assimilation (instead of L-band brightness). The use of the arithmetic mean over the estimation pixels within each observation pixel becomes questionable at higher microwave frequencies.

   (b) The rainfall interception model proved very difficult to include in the assimilation algorithm. This is not surprising given the disparity in the time scales of canopy interception and of the other land surface states. Since the quantity of water stored in the canopy is very small, the influence of the interception process on the soil moisture estimates is limited. On the other hand, canopy interception provides invaluable information about brightness data. Due to complicated reflections at and within the canopy, brightness images are essentially useless when the canopy is wet.

   Simply neglecting canopy interception altogether, as we have done so far, is one way out of the dilemma, but this works only when the brightness images are subject to careful quality control. In future implementations, it might be beneficial to run the interception model off-line and in forward mode only. Brightness images taken at times when the canopy is wet can then be discarded in a quality control step.

(c) For continental-scale applications, our land surface model will have to be re-evaluated. The physics of the model may not be appropriate at much larger scales, although no obvious alternatives spring to mind. In addition, more computational savings may be required before a continental-scale application becomes feasible.

(d) It is desirable to couple the land surface model with a runoff and a groundwater model in order to more accurately represent land surface hydrologic processes.

(e) Any change in the land surface model must also be reflected in its adjoint.

7. Since the iterated indirect representer algorithm is the cheapest optimal approach, it constitutes an invaluable benchmark. The actual computational requirements, however, may defy its use for operational applications.

   (a) The results on the computational demands and the scalability of the algorithm (Chapter 8) must be corroborated or modified through numerical experiments conducted with more pixels and longer experiment periods.

   (b) With its CPU and memory requirements, the iterated indirect representer method may be too expensive for an operational assimilation package for some time still. Moreover, there is little hope of calculating accurate and detailed posterior error covariances even in research studies.

   (c) Nonetheless, the iterated indirect representer approach is an efficient way to derive the optimal estimates, something which is not possible with a Kalman filter or an adjoint-based gradient search at the same computational expense. The representer method can therefore be used as a benchmark against which cheaper but less optimal algorithms can be evaluated.